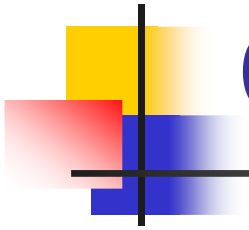
# Information Theoretic Model for Inference Resistant Knowledge Management in RBAC Based Collaborative Environment

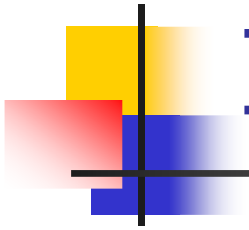
Manish Gupta

Sivakumar Chennuru



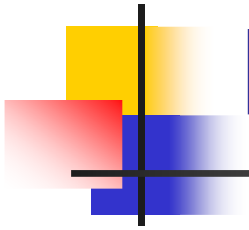
# Overview

- Model to reveal inference vulnerabilities
- Need for the model
- Description of the model
- Results
- Benefits



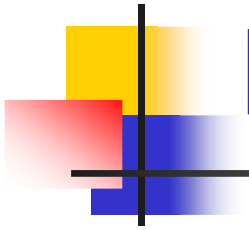
# Introduction

- Information – key organizational resource
  - Dissemination and Sharing
- Current Access Control Methods
  - Segregation techniques
  - Direct Access Control
- Are these sufficient?



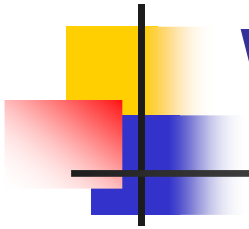
# Need for the model

- Indirect Access Mechanisms
  - Individual knowledge
  - On the role knowledge acquisition
  - Informal communication channels
- Framework for identifying and analyzing
  - Data (information)
  - Roles
  - Roles' direct access to data
  - Association among roles prone to inference



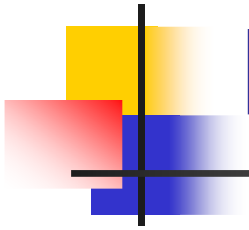
# Prior work

- Database design
  - Uncover secondary paths leading to inferences
  - Functional dependencies
  - Conceptual structures
  - Semantic data modeling



# Why Information Theory?

- Mathematical theory to quantify the concept of information
- Measure for the Entropy and Information
- Mutual information
  - Amount of information obtained by observing another information
- Channel
  - Interaction between employees with different roles
  - Continuous transfer over a variable length of time



# Model Description

- Data Units
  - ORG = $\{D_1, D_2, \ldots\ldots, D_N\}$; where N is total number of data units in the organization.
  - Each data unit $D_i$ will have some information content $I_i$
  - Each data unit may or may not be linked with other data units.
  - The information revealed is additive if the data units are statistically independent.

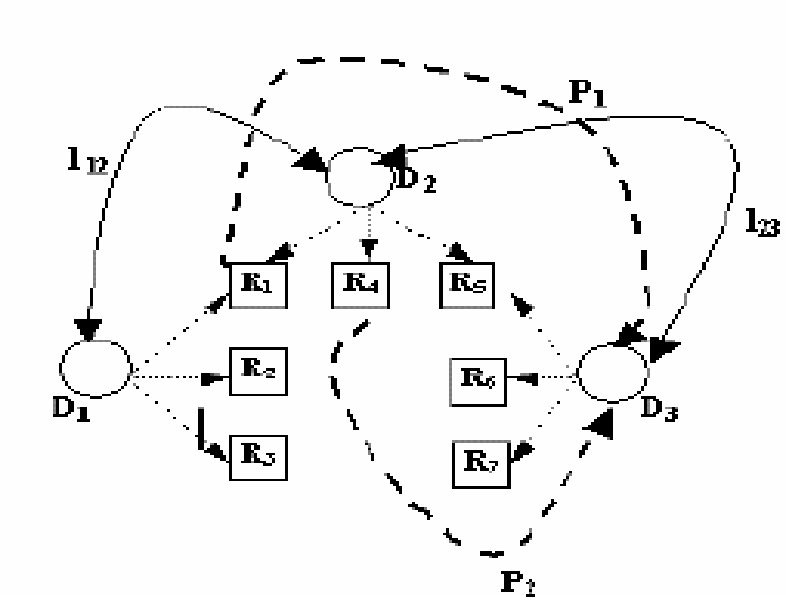
# Model Description

- Data Units (contd)
  - The mutual information of a data unit I(i;j) is the difference in the uncertainty of $D_i$ and the remaining uncertainty of $D_i$ after observing $D_j$ .
- Data Inputs
  - For each data unit Di , all data units in set ORG which are not statistically independent
  - For each data unit Di , all proper subsets of ORG which are not statistically independent

# Model Description

- Roles
  - Set of Roles in the organization, R = { $R_1$, $R_2$,....., $R_M$}; where M is total number of roles in the organization
  - Relationship between Data units and Roles
  - Relationship between Roles
  - Degree of Proximity of Roles

# Roles and Data Units




- Relation between roles

  RLINK1 [$R_1$] = { $R_2$ , $R_3$, $R_4$, $R_5$ }

  RLINK2 [$R_1$] = { $R_6$ , $R_7$ }

- Role-Data unit direct access

  RSET [$D_1$] □ { $R_1$ , $R_2$ , $R_3$ }

  RSET [$D_2$] □ { $R_1$ , $R_4$ , $R_5$ }

  RSET [$D_3$] □ { $R_5$ , $R_6$ , $R_7$ }

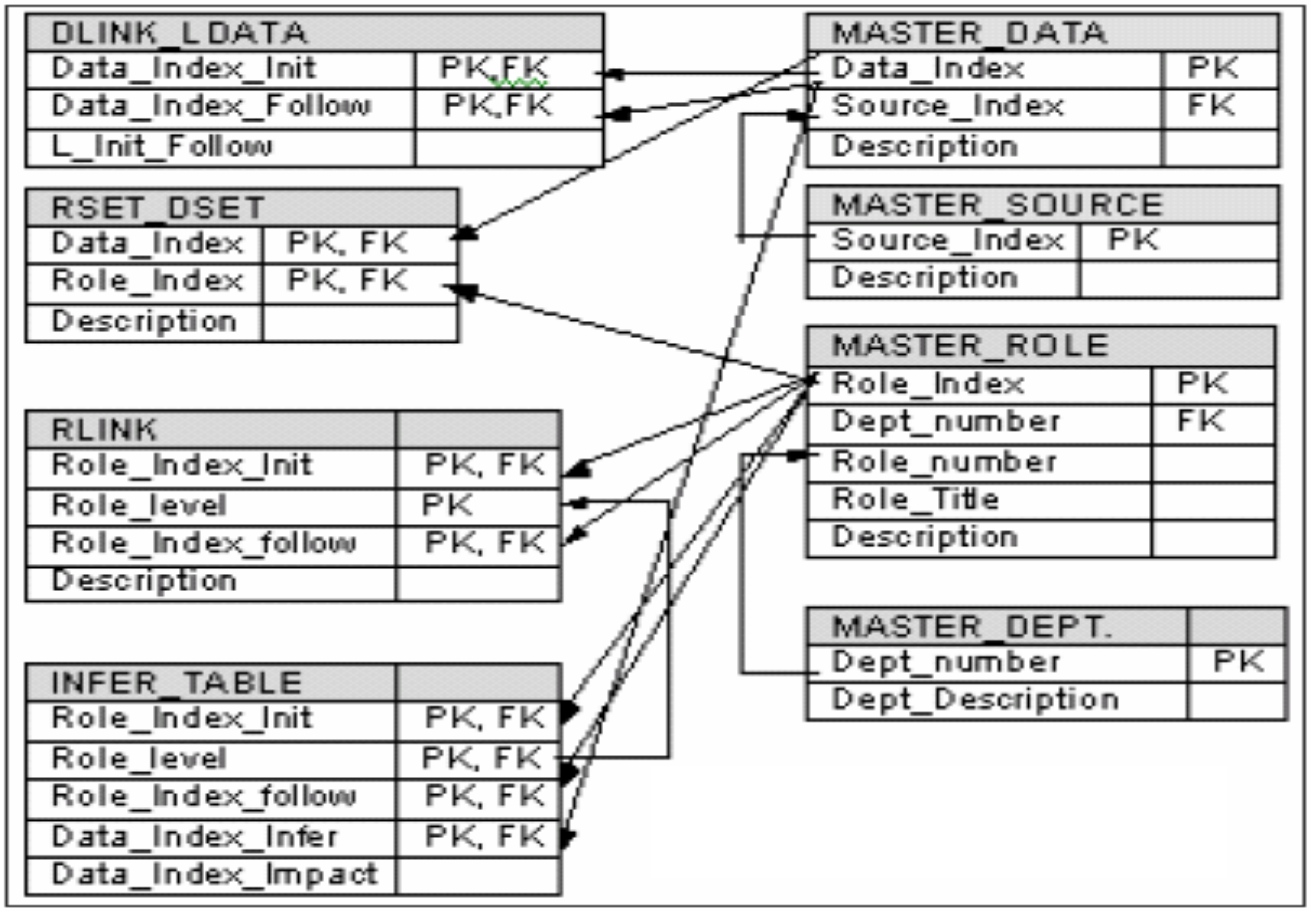
- Role-Data unit Indirect Access

  $R_4 \longrightarrow D_3$ (path $P_2$ )

  $R_1 \longrightarrow D_3$ (path $P_1$ )

  Strength of inference depends upon mutual information.

# Proposed ER Model



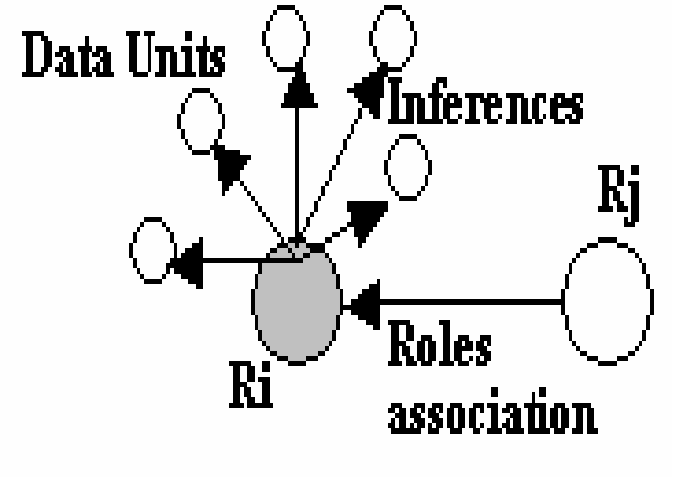
# Inference Extraction

- Select a role (r) from MASTER_ROLE
- Select all the data units ($d_i$) linked to the above role from RSET_DSET
- Select all the roles linked to the above role from RLINK
- Select the data units ($d_k$) accessed by the linked roles and the mutual information of these data units ($d_k$;$d_i$) from data units accessed by the role (r)
- The results are stored in INFER_TABLE

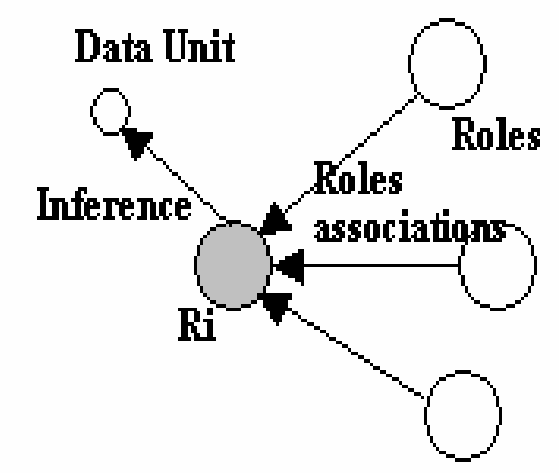
# Results

- Role centric views

  Roles and Role associations that can be exploited for inference attacks.




Scenario 1

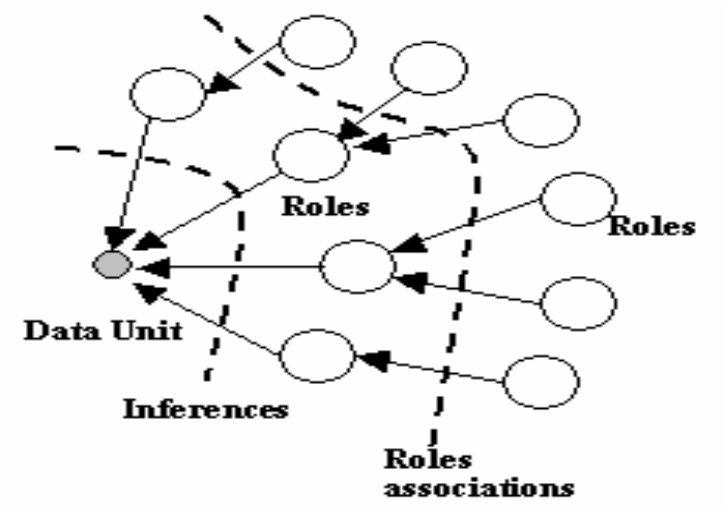


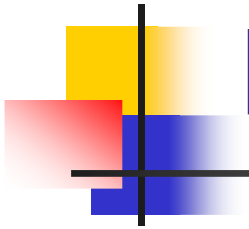

Scenario2

# Results

- Data centric views

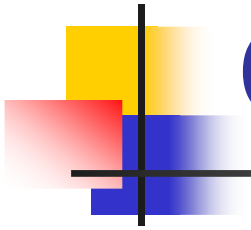
  List of data units most vulnerable to design with the given role structure.





# Benefits

- Identifying possible inference attacks
- Assignment of individuals to the roles
- Greater assurance against insider attacks



# Questions

- Thank You